An overview of the OAIS and Representation Information

JORUM, DCC and JISC Forum
Long-term Curation and Preservation of Learning Objects
February 9th 2006
University of Glasgow

Manjula Patel
UKOLN and DCC
University of Bath, UK

Funded by: JISC e-Science dti

Presentation Outline

- The OAIS Reference Model
  - Background
  - Concepts
  - Functional Model
  - Information Model
  - Representation Information
- DCC Development
  - Agenda
  - DCC Representation Information Registry
  - DCC-RIR: a basis for collaboration

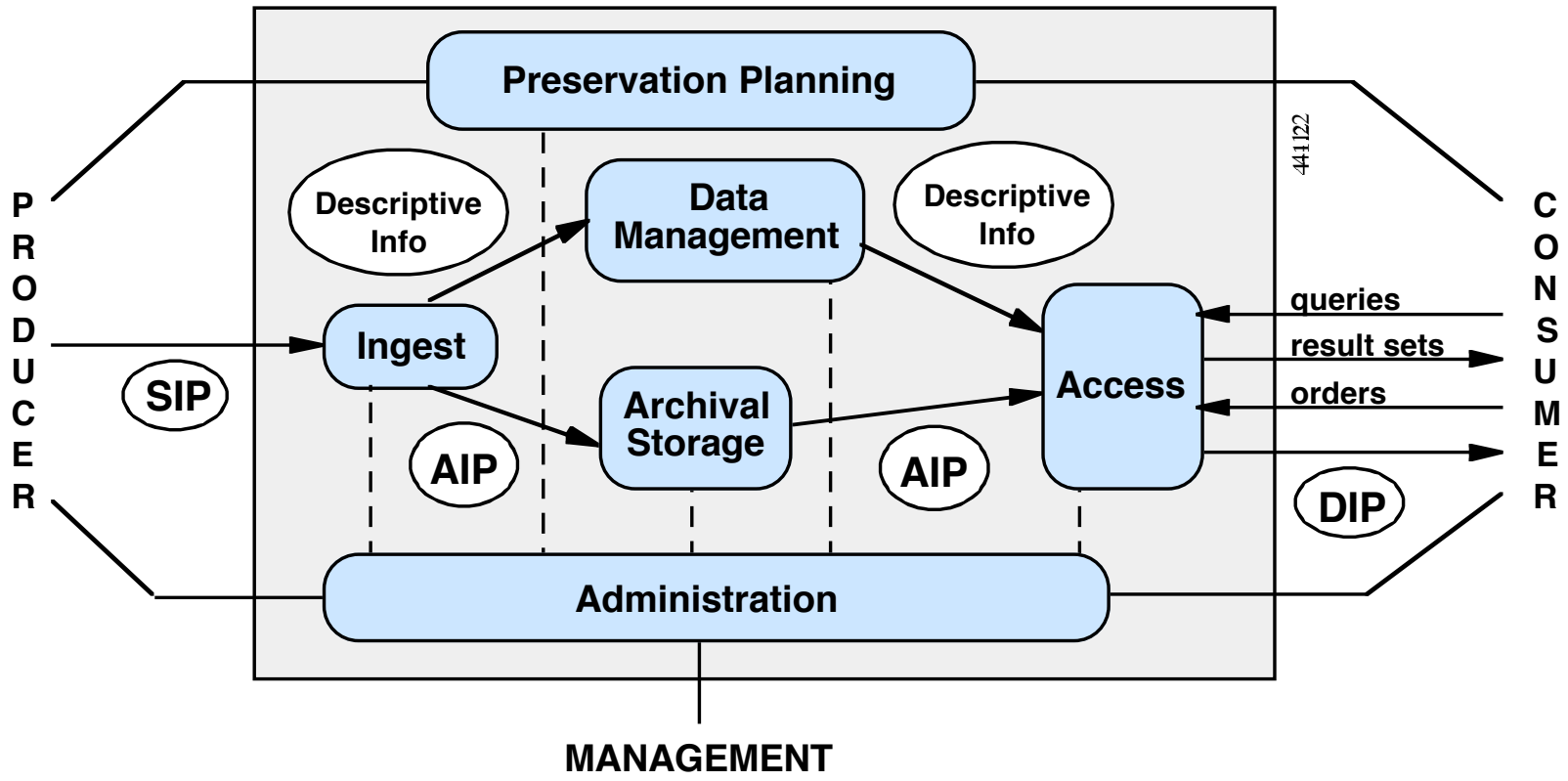
OAIS Background

- OAIS -*Reference Model for an Open Archival Information System*
  http://www.ccsds.org/documents/650x0b1.pdf
- Development led by the Consultative Committee for Space Data Systems (CCSDS)
- Adopted as ISO 14721:2003
- "Open" refers to development of the model in an open forum
- Reference Model, not a blueprint for implementation
- Establishes a common framework of terms and concepts
- Identifies the basic functions of an OAIS
- Defines an information model
- Three major areas of influence:
  - Preservation metadata schemas
  - Architecture and system design
  - Conformance criteria for archival repositories

OAIS Definition and Selected Concepts

- OAIS: "An archive, consisting of an organization of people and systems, that has accepted the responsibility to preserve information and make it available for a Designated Community"
- Designated Community: Community of stakeholders and users that the OAIS serves
- Knowledge Base: A set of information, incorporated by a user or system, that allows that user or system to understand the received information
- Information Object: Data Object + Representation Information
- Representation Information: **any** information required to render, interpret and understand digital data
- Information Package: Content Information + Preservation Description Information + Packaging Information (Submission, Archival and Dissemination Information Packages)
- Preservation Description Information: Provenance, Context, Reference, Fixity information

OAIS Functional Model

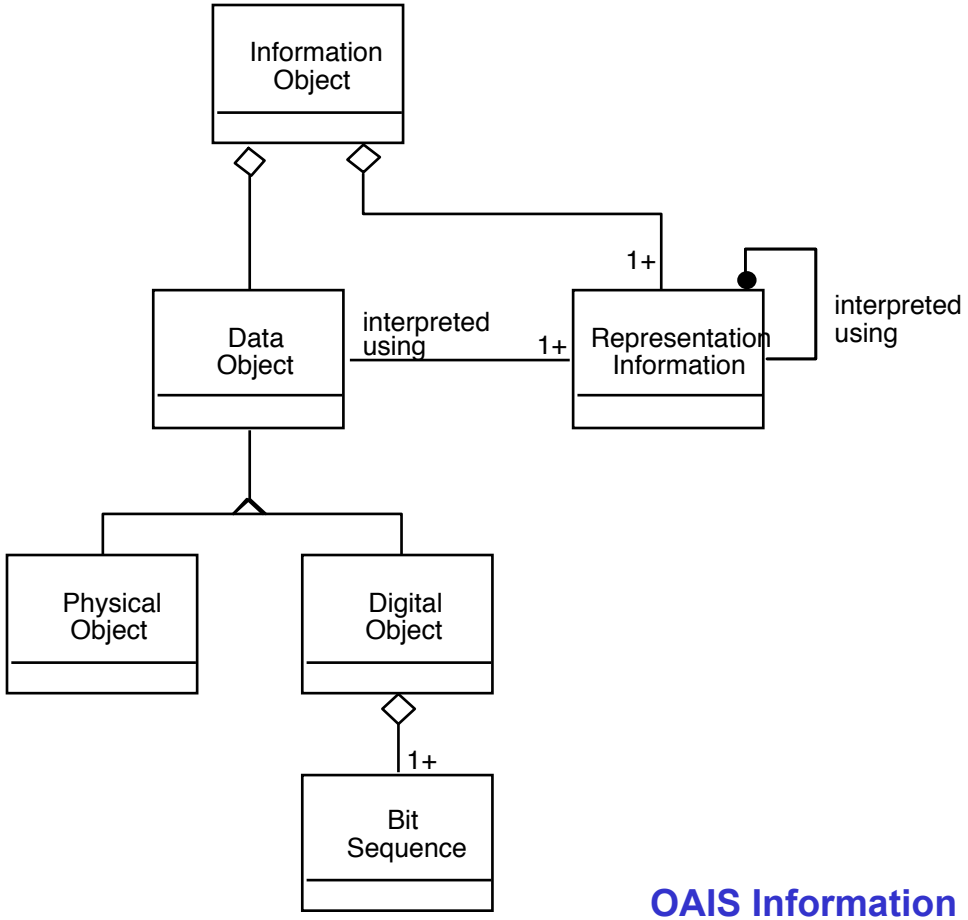


OAIS Functional Entities (Figure 4-1)

OAIS Functional Entities

- Ingest: services and functions that accept SIPs from Producers; prepares AIPs for storage, and ensures that AIPs and their supporting Descriptive Information become established within the OAIS
- Archival Storage: services and functions used for the storage and retrieval of AIPs
- Data Management: services and functions for populating, maintaining, and accessing a wide variety of information
- Administration: services and functions needed to control the operation of the other OAIS functional entities on a day-to-day basis
- Preservation Planning: services and functions for monitoring the OAIS environment and ensuring that content remains accessible to the Designated Community
- Access: services and functions which make the archival information holdings and related services visible to Consumers

OAIS Information Object



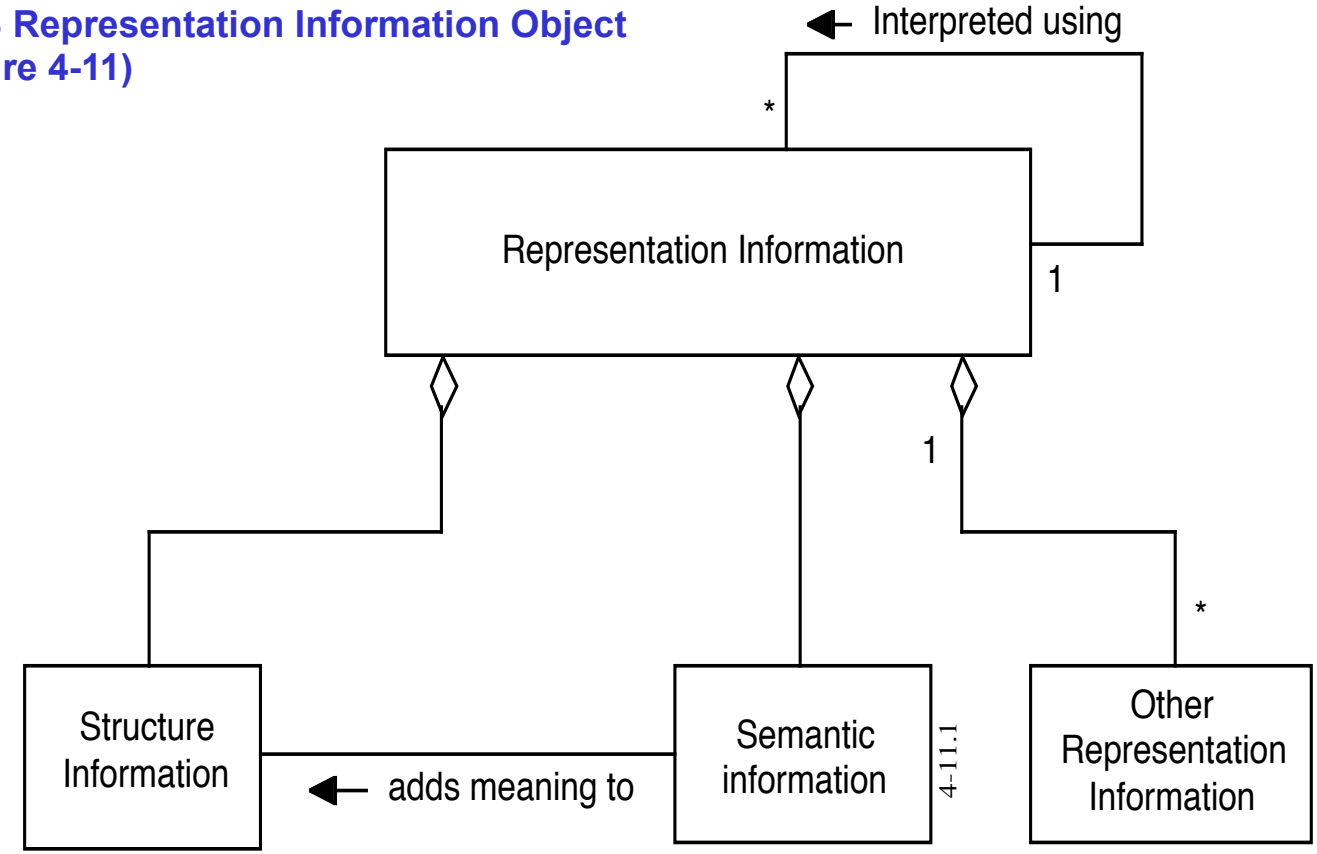
OAIS Information Object (Figure 4-10)

OAIS Representation Information

- Representation Information: **any** information required to render, interpret and understand digital data (includes file formats, software, algorithms, standards, semantic information etc.)
- Representation Information is recursive in nature
- E-materials such as Learning Objects are typically
  - Complex (compound, comprising many individual assets)
  - Mixed media (text, graphics, simulations etc.)
  - Multimedia (video, audio, interactive etc.)
- Essential that Representation Information itself is curated and preserved to maintain access to (render and interpret) digital data

OAIS Representation Information Network

OAIS Representation Information Object (Figure 4-11)



DCC Development

"DCC Approach to Digital Curation" white paper sets out the path for development activities based on OAIS,
http://dev.dcc.ac.uk/twiki/bin/view/Main/DCCApproachToCuration

- Monitoring international standards
- Development of a Representation Information registry/ repository (DCC-RIR)
- Recommendations for tools and methods for generating Representation Information
- Creating test-beds for digital curation tools
- Creating auditing and certification processes for trusted repositories

DCC-RIR

- Representation Information (RI) is the key to long-term access
- DCC-RIR should be OAIS compliant
- Emphasis on interoperability and automated use
- Vision is to have a global, distributed network of RI
- Provide an infrastructure of reliable and trusted RI which other archives can rely on
- Investigate how RI fits into the work of other projects and initiatives

DCC-RIR: A basis for collaboration

- Idea of RI is the key
  - Information Object: a specific object to be archived
  - RI: all information required to interpret and render the object
  - RI Label: used to connect RI to an Information Object
- RI label serves as a mechanism for accessing RI in the DCC-RIR
  - A label attached to each digital object
  - Label should identify RI (persistent identifier)
  - Provides mechanism for combining individual RI components
  - May be a structured digital object itself (to cope with packaging of multiple objects)
- Need to see how RI and RI label might fit into work of other projects and architectures

OAIS Responsibilities and Conformance

- OAIS Mandatory Responsibilities:
  - Negotiating and accepting information
  - Obtaining sufficient control of the information to ensure long-term preservation
  - Determining the "designated community"
  - Ensuring that information is "independently understandable"
  - Following documented policies and procedures
  - Making the preserved information available

- Many repositories or preservation tools claim OAIS compliance:
  - e.g., DSpace, OCLC Digital Archive, METS, LOCKSS etc.

OAIS…More

- Conformance and Certification
  - OCLC/RLG Digital Archive Attributes Working Group (Report on Trusted Digital repositories, 2002)
  - RLG-NARA Task Force on Digital Repository Certification (Draft checklist for self-certification, August 2005)
- Archival Information Units and Archival Information Collections
- Information Package transformations, e.g. for Ingest and Access
- Preservation perspectives:
  - Migration e.g refreshment, replication, repackaging, transformation
  - Preservation of look and feel (e.g. emulation, virtual machines)
- Archive interoperability, e.g. P2P, federation

References

- OAIS Reference Model: http://www.ccsds.org/documents/650x0b1.pdf
- DPC Technology Watch Report on OAIS model by Brian Lavoie (OCLC Research): http://www.dpconline.org/
- RLG/NARA Task Force on Digital Repository Certification: http://www.rlg.org/
- DCC Development White Paper “DCC Approach to Digital Curation under Development”: http://dev.dcc.ac.uk/twiki/bin/view/Main/DCCApproachToCuration

Thank you for your attention!

Acknowledgements
Several slides adapted from those by Michael Day

Manjula Patel
UKOLN, University of Bath, UK
m.patel@ukoln.ac.uk
http://www.ukoln.ac.uk/
http://www.dcc.ac.uk/